REPORT DOCUMENTATION PAGE

AFRL-SR-AR-TR-03-

0307

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
		1 Feb 00 - 30 Nov 02

4. TITLE AND SUBTITLE
Competent Probabilistic Model Building Genetic Algorithms

5. FUNDING NUMBERS
F49620-00-1-0163

6. AUTHOR(S)
David Goldberg

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Department of General Engineering
The University of Illinois at Urbana-Champaign
117 Transportation Building
104 S. Mathews Avenue
Urbana, IL 61801

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AFOSR/NM
4015 Wilson Blvd, Room 713
Arlington, VA 22203-1954

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
F49620-00-1-0163

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION AVAILABILITY STATEMENT
APPROVED FOR PUBLIC RELEASE, DISTRIBUTION UNLIMITED

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*
AFOSR Grant No. F49620-00-1-0163 has come to a close with substantial progress on all fronts. In particular, the Bayesian Optimization Algorithm has been extended to the important class of hierachically difficult problems. Progress has also been made in a number of related competent GAs and efficiency enhancement techniques as well as in the application of these techniques.

20030818 014

14. SUBJECT TERMS			15. NUMBER OF PAGES
			16
			16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT

Progress and Final Report
Competent Probabilistic Model Building
Genetic Algorithms
AFOSR Grant No. F49620-00-1-0163
01 December 1999 to 30 November 2002

David E. Goldberg
Department of General Engineering
IlliGAL Report No. X0016 (Limited Circulation)
July 2003

Department of General Engineering
University of Illinois at Urbana-Champaign
117 Transportation Building
104 South Mathews Avenue
Urbana, IL 61801

Progress & Final Report for
Competent Probabilistic Model Building Genetic Algorithms
AFOSR Grant No. F49620-00-1-0163
1 December 1999 to 30 November 2002

David E. Goldberg
Department of General Engineering
The University of Illinois at Urbana-Champaign
117 Transportation Building
104 S. Mathews Avenue
Urbana, IL 61801

Abstract

AFOSR Grant No. F49620-00-1-01630 (Competent Probabilistic Model Building Genetic Algorithms) has come to a close with substantial progress on all fronts. In particular, the Bayesian Optimization Algorithm has been extended to the important class of hierarchically difficult problems. Progress has also been made in a number of related competent GAs and efficiency enhancement techniques as well as in the application of these techniques.

1 Objectives

The objectives of AFOSR Grant No. F49620-00-1-01630 (Competent Probabilistic Model Building Genetic Algorithms) were as follows:

1. Develop, implement, and enhance two probabilistic model building GAs: BOA (Bayesian optimization algorithm) and ecGA (extended compact GA).

2. Test BOA and ecGA on rigorous theoretical function test suite.

3. Extend BOA and/or ecGA to hierarchically difficult problems.

4. Extend BOA and/or ecGA to alternative codings.

5. Apply the developed algorithms to two problems of Air Force interest.

Significant progress has been made on each fronts. This progress is briefly reviewed in the next section.

2 Status of Effort

The project has come to a close with key progress in each of the objective areas above. Key areas of progress are highlighted below.

hBOA. A major task of the project was to advance the cause of *probabilistic model-building genetic algorithms*, a particular type of *competent* GA that solve hard problems, quickly, reliably, and accurately. Progress in this area over the course of the project has been substantial and included particular concentration on the hierarchical Bayesian optimization algorithm (hBOA). Martin Pelikan's groundbreaking dissertation on hBOA is arguably the most notable achievement of the project. Recall that simple BOA builds a Bayesian network model of the best points in each generation, using this model in place of crossover, mutation, and other genetic operators to speculate on the best points in the search space. Because of the generality of the Bayes net assumption, BOA can quickly search for optimal solutions under more general conditions than can a linkage-oriented solver. Hierarchical BOA takes the original scheme, combines it with *niched* selection and *chunking* techniques to form a technique that quickly solves hierarchically difficult problems.

The technique has been theoretically and empirically shown to be sub- or nearly quadratic over a very large class of problems—to what Herbert Simon called the class of *nearly decomposable* problems. This effort has resulted in a the filing of a patent through the University of Illinois Office of Technology Management.

hBOA on maxSAT and spin glasses. We have published work on applying hBOA to satisfiability problems and nearest neighbor spin glasses in two and three dimensions. Of course, worst case satisfiability problems are intractable, but if substructure exists, hBOA appears to be able to detect it and exploit it. In spin-glass problems, hBOA is competitive with problem-specific heuristics.

Multiobjective hBOA. Nazan Khan wrote a useful MS thesis extending BOA and hBOA to difficult multiobjective problems. Simply stated, a standard selection technique of evolutionary multiobjective optimization (EMO), NSGA-2, was added to Pelikan's hBOA, thereby creating a very powerful multiobjective solver.

Hybrid BOA and ES. First steps were taken to use BOA in real-coded problems by using a discrete BOA solver in concert with self-adaptive evolution strategies. The approach was successful in combining the global search capability of BOA with the local rapidity of an ES.

Other model builders invented. Other model-building techiques were invented as part of this effort. In particular, standard techniques of machine learning such as ID3 have been adapted to the task (Llora & Goldberg, 2003) as have techniques inspired by design structure matrices or DSMs (Yu, Yassine, & Goldberg, 2003). The latter effort resulted in a patent disclosure for both the GA and organizational implications of our work,

Knjazew's OmeGA. Dimitri Knjazew, a visiting scholar from Dortmund, Germany, and the PI developed OmeGA, the ordering messy genetic algorithm for solving hard problems over permutations. Using previously supported fast messy GA technology, the researchers have come up with a way to solve difficult scheduling and sequencing problems quickly, reliably, and accurately. This work is being published as a book by Kluwer Academic Publishers.

Tsutsui's real-Coded linkage GA & PMBGA. Professor Shigeyoshi Tsutsui finished his visit from Hannan University and his work on two new techniques has been published. Tsutsui and

Pelikan worked together to carry the framework of BOA over to real-coded problems with success.

LLGA revisted. Earlier work in the lab developed the so-called *linkage learning* GA. Y. P. Chen has revisited that work and extended it by using encapsulating punctuation marks to improve building block separation and mixing. A previous defect of the LLGA was its inability to solve problems with multiple uniformly scaled building blocks simultaneously. To carry this work forward, new theory has been developed of *tightness time* to predict when good linkage will be achieved.

Advanced scalability theory. Mixing is arguably the crucial step to achieve scalable GAs. Sastry's work (Sastry & Goldberg, 2003) advances the state of the art by building principled models of population size and run duration for one-point and other low-mixing crossover operators.

Time utilization. Ravi Srivastava completed an MS thesis on the theory of time continuation and its validation. This work settles the great debate between crossover and mutation on economic and solution-quality grounds. Essentially, the theory shows that a combination of the serial processing of mutation and the parallel processing of mixing or crossover is optimal in most problems and the degree of balance is related to the salience of the building blocks or other features of the problem that give it a sequential nature.

Evaluation relaxation. Theses by Laura Albert and Kumara Sastry add to our knowledge about time-quality tradeoffs in fitness evaluation. Sastry's work contrasts the temporal strategies needed in problems of pure *bias* to spatial strategies required in problems of pure *variance*. Albert's work applies these ideas to evaluation relaxation in problems based on the numerical solution of integral equations. The key result is that at any given point in a run, the accuracy of the numerical integration should be of the same order as the precision required of each decision variable, and discretization is *scheduled* in a principled manner to make this happen.

Hybridization. A first analysis of the optimization of global-local search was performed in 1999 and Abhishek Sinha has completed an MS thesis on this topic. The initial theory required static probabilities of hitting a basin of attraction to the local search. With an adaptive global search technique like a GA, those probabilities must be assumed to vary. Prasanna Parthasarathy has studied hybrid structural optimization and has developed a novel technique of Baldwinian sharing for such problems as well as a technique to save function evaluations by identifying basins of attraction approximately. Practical self-adaptive techniques were studied in concert with Minsker and her PhD student Espinoza.

Parallelization. Cantu-Paz's work in 1999 continues to inform our understanding of GA parallelism in a sufficiently adequate and accurate manner. A number of items were revisited as a result of this study. The issue of mult- versus single-run studies was analyzed, and multiple independent runs are sometimes advantageous. Master-slave parallelism was revisited when significant GA overhead exists. The standard formulae for optimal slave counts remain valid, although overall expressions have been derived that show the additional degradation to speedup in cases with significant GA overhead.

Time-quality analysis of different representations. Franz Rothlauf completed his dissertation investigating code redundancy, locality, and scaling. His framework closes the loop on coding design by investigating codes in a quality-time framework for the first time. This is a critical milestone and will help put the study of representations on a more scientific footing that has been possible

until now.

Symmetry and multimodality studied. Clarissa Van Hoyweghen has completed her study of spin glasses and certain network problems with massive and essential symmetry and multimodality. Walsh analyses have been performed and these have resulted in basic understanding of the challenges these problems pose for genetic algorithms. The problems studied to date have succumbed to GAs augmented with effective nichers such as sharing, restricted tournament selection, deterministic crowding, and the like.

Application of GAs and GP in material science. Working with Duane Johnson here at Illinois, Kumara Sastry has used GAs and GP to speed the modeling of advanced materials 9 orders of magnitude. This work is continuing.

Electromagnetic applications underway. Contact has been made with researchers at Hanscom AFB and studies are currently underway using Matlab code written at Hanscom as well as standard EM analysis procedures (NEC).

Books completed and published. Cantu-Paz's *Efficient and Accurate Parallel Genetic Algorithms* (Kluwer) was published in 1999. Butz's, *Anticipatory Classifier Systems* (Kluwer) was published in 2002. Knjazew's *OmeGA* (Kluwer) was published in 2002. The PI's *Design of Innovation* (Kluwer) was published in 2002. Rothlauf's *Representations for Genetic and Evolutionary Algorithms* (Springer) was published in 2002. Pelikan is publishing a volume from his hBOA dissertation with Springer.

3 Personnel Supported

This section details the individuals supported on these projects.

3.1 Faculty supported

Professor David E. Goldberg, the principal investigator, was supported during the summers of 2000–2002.

3.2 Faculty or post-docs supported by or affiliated with the project

The following is a list of visiting faculty supported by or affiliated with the project. Unsupported affilates may have some travel or incidental expenses paid by the project.

1. Professor Matthieu Domaszewski (Universite de Technologie de Belfort-Montbeliard, supported partially by AFOSR)

2. Professor Felipe Padilla (University Autonomous of Aguascalientes , Mexico)

3. Professor T. K. Kurian (Calicut Regional Engineering College, India, supported partially by AFOSR)

4. Professor Shigeyoshi Tsutsui (Hannan University, Japan)

5. Professor Pier Luca (Politecnico di Milano, Italy)

6. Dr. Xavier Llora (Ramon Llull University, Barcelona, Spain)

7. Dr. Kei Ohnishi (Kyushu University, Japan)

3.3 Graduate students supported & affiliated with project

The following is a list of graduate students supported or affiliated with the project Unsupported affilates may have had travel or incidental expenses supported by the project.

1. L. Albert (UIUC, supported by AFOSR)

2. M. Butz (University of Wuerzburg, Germany & NCSA)

3. Y. P. Chen (UIUC, supported on TA)

4. A. Ceroni (University of Florence, Italy)

5. N. Khan (UIUC, NCSA)

6. A. Kosorukoff (UIUC, supported by UI Research Board & TEC)

7. P. Parthasarathy (UIUC, supported by AFOSR & NSF)

8. M. Pelikan (UIUC, supported by AFOSR)

9. F. Rothlauf (University of Bayreuth, Germany)

10. K. Sastry (UIUC, supported by AFOSR & NSF)

11. A. Sinha (UIUC, supported by AFOSR)

12. R. Srivastava (UIUC, supported by NSF)

13. T.-L. Yu (UIUC, supported by AFOSR)

14. K. Sastry (UIUC, supported by AFOSR, Army, & NSF)

15. C. Van Hoyweghen (University of Antwerp, Belgium)

3.4 Undergraduate students supported by or affiliated with the project

The following is a list of undergraduate research assistants who have been supported by or affiliated with the project:

1. L. Babayeva (AFOSR)

2. J. Borgerson (AFOSR & NSF)

3. C. Cunningham (NSF)

4. J. Daly (NSF)

5. K. Hawley (AFOSR & NSF)

6. K. Kakugawa (NSF REU researcher)

7. J. Leesman (AFOSR & NSF)

8. T. Latoza (NSF REU researcher)

9. M. Lipinski (AFOSR)

10. M. Perry (AFOSR)

11. N. Sis (AFOSR)

12. A. Vaughn (NSF

13. J. Wilson (AFOSR)

4 Publications

Publications submitted, accepted, and published over the course of the project are presented in this section.

4.1 Submitted

Reed, P., Minsker, B. S., & Goldberg, D. E. (2002). Simplifying multiobjective optimization: An automated design methodology for the nondominated sorted genetic algorithm-II. *Water Resources Research.*

Sastry, K., Johnson, D. D., Goldberg, D. E., & Bellon, P. (2003). Genetic programming for multi-timescale modeling: A proof-of-principle study. *Physical Review Letters.*

4.2 Accepted, Not Yet Published

Butz, M. V., & Goldberg, D. E. (in press). Generalized state values in an anticipatory learning classifier system. In M. V. Butz, Sigaud, O., & Gerard, P. (Eds.), *Anticipatory behavior in adaptive learning systems.* Heidelberg, Germany: Physica-Verlag.

Butz, M., & Goldberg, D. E. (in press). Bounding the population size in XCS to ensure reproductive opportunities. *GECCO-2003: Proceedings of the Genetic and Evolutionary Computation Conference.*

Butz, M., & Goldberg, D. E. (in press). Tournament selection: Stable fitness pressure in XCS. *GECCO-2003: Proceedings of the Genetic and Evolutionary Computation Conference.*

Butz, M., Goldberg, D. E., & Tharakunnel, K. (in press). Analysis and improvement of fitness exploitation in XCS: Bounding models, tournament sselection, and bilateral accuracy. *Evolutionary Computation.*

Cantu-Paz, E. & Goldberg, D. E. (in press). Are multiple runs of genetic algorithms better than one. *GECCO-2003: Proceedings of the Genetic and Evolutionary Computation Conference.*

Chen, Y.-P., & Goldberg, D. E. (in press). Tightness time for the linkage learning genetic algorithm. *GECCO-2003: Proceedings of the Genetic and Evolutionary Computation Conference.*

Chen, Y.-P., & Goldberg, D. E. (in press). An analysis of a reordering operator with tournament selectio on a GA-hard problem. *GECCO-2003: Proceedings of the Genetic and Evolutionary Computation Conference.*

Espinoza, F., Minsker, B. & Goldberg, D. E. (in press). Performance evaluation and population reduction for a self-adaptive hybrid genetic algorithm (SAHGA). *GECCO-2003: Proceedings of the Genetic and Evolutionary Computation Conference.*

Goldberg, D. E. (in press). John H. Holland, facetwise models, and economy of thought. In L. Booker, S. Forrest, M. Mitchell, & R. Riolo (Eds.), *Perspectives on Adaptation in Natural and Artificial Systems.* New York: Oxford University Press.

Llora, X., Goldberg, D. E., Traus, I., & Bernado, E. (in press). Accuracy, parsimony, and generality in evolutionary learning systems via multiobjective selection. *Advances in Learning Classifier Systems.*

Llora, X., & Goldberg, D. E. (in press). Bounding the effect of noise in multiobjective learning classifier systems. *Evolutionary Computation.*

Llora, X., & Goldberg, D. E. (in press). Wise breeding GA via machine learning techniques for function optimization. *GECCO-2003: Proceedings of the Genetic and Evolutionary Computation Conference.*

Pelikan, M. (in press). *Hierarchical Bayesian optmization algorithm: Toward a new generation of evolutionary algorithms.* Heidelberg, Germany: Physica-Verlag.

Pelikan, M., & Goldberg, D. E. (in press). Hierarchical BOA solves Ising spring glasses and MAXSAT. *GECCO-2003: Proceedings of the Genetic and Evolutionary Computation Conference.*

Sastry, K., & Goldberg, D. E. (in press). Scalability of selectorecombinative genetic algorithms for problems with tight linkage. *GECCO-2003: Proceedings of the Genetic and Evolutionary Computation Conference.*

Tharakunnel, K, Butz, M., & Goldberg, D. E. (in press). Towards building block propagation in XCS: A negative result and its implications. *GECCO-2003: Proceedings of the Genetic and Evolutionary Computation Conference.*

Yu, T.-L., Goldberg, D. E., Yassine, A. (in press). A genetic algorithm design inspired by organizational theory. *GECCO-2003: Proceedings of the Genetic and Evolutionary Computation Conference.*

Yu, T.-L., Goldberg, D. E., & Sastry, K. (in press). Optimal sampling and speed-up for genetic algorithms on the sample OneMax problem. *GECCO-2003: Proceedings of the Genetic and Evolutionary Computation Conference.GECCO-2003: Proceedings of the Genetic and Evolutionary Computation Conference.*

4.3 Published

Butz, M. V., Goldberg, D. E., & Stolzmann, W. (2000). Probabililty-enhanced predictions in the anticipatory classifier system. *Proceedings of the 2000 International Workshop on Learning Classifier Systems.*

Butz, M. V., Goldberg, D. E., & Stolzmann, W. (2000). Investigating generalization in the anticipatory classifier system. *Proceedings of the 2000 Parallel Problem Solving from Nature Conference, 735–744.*

7

Butz, M. V., Goldberg, D. E., & Stolzmann, W. (2000). Introducing a genetic generalization pressure to the anticipatory classifier system, Part I: Theoretical approach. *GECCO-2000 Proceedings of the Genetic and Evolutionary Computation Conference*, 34–41.

Butz, M. V., Goldberg, D. E., & Stolzmann, W. (2000). Introducing a genetic generalization pressure to the anticipatory classifier system, Part 2: Performance analysis. *GECCO-2000 Proceedings of the Genetic and Evolutionary Computation Conference*, 42–49.

Cantú-Paz, E. (2000). Efficient and accurate parallel genetic algorithms. Boston, MA: Kluwer Academic Publishers.

Cantú-Paz, E., & Goldberg, D. E. (2000). Efficient parallel genetic algorithms: Theory and practice. *Computational Methods in Applied Mechanics and Engineering, 186*, 221-238.

Goldberg, D. E. (2000). The design of innovation: Lessons from genetic algorithms, lessons for the real world. *Technological Forecasting & Social Change, 64*, 7–12.

Goldberg, D. E., & Deb, K. (2000). Genetic and evolutionary computation in computational mechanics and engineering: Preface to a special edition. *Computational Methods in Applied Mechanics and Engineering, 186*, 121-124.

Goldberg, D. E. (2000). A hydroinformatician's approach to computational innovation and the design of genetic algorithms. *Journal of Hydroinformatics, 2*(3), 155–162.

Harik, G., & Goldberg, D. E. (2000). Linkage learning through probabilistic expression. *Computational Methods in Applied Mechanics and Engineering, 186*, 295–310.

Knjazew, D. & Goldberg, D. E. (2000). Large-scale permutation optimization with the ordering messy genetic algorithm. *Proceedings of the 2000 Parallel Problem Solving from Nature Conference*, 631–640.

Knjazew, D. & Goldberg, D. E. (2000). OMEGA—Ordering Messy GA: Solving permutation problems with the fast lmessy genetic algorithm and random keys. *GECCO-2000 Proceedings of the Genetic and Evolutionary Computation Conference*, 181–188.

Pelikan, M. & Goldberg, D. E. (2000). Genetic algorithms, clustering, and the breaking of symmetry. *Proceedings of the 2000 Parallel Problem Solving from Nature Conference*, 385–394.

Pelikan, M. & Goldberg, D. E. (2000). Hierarchical problem solving by the Bayesian optimization algorithm. *GECCO-2000 Proceedings of the Genetic and Evolutionary Computation Conference*, .

Pelikan, M., Goldberg, D. E., & Cantú-Paz, E. (2000). Bayesian optimization algorithm, population sizing, and time to convergence. *GECCO-2000 Proceedings of the Genetic and Evolutionary Computation Conference*, 275–282.

Pelikan, M., Goldberg, D. E., & Cantú-Paz, E. (2000). BOA: The Bayesian Optimization Algorithm. *Evolutionary Computation*, 267–274.

Pelikan, M., Goldberg, D. E., & Lobo, F. (2000). A survey of optimization by building and using probabilistic models. *Proceedings of the 2000 American Controls Conference, 5*. 3289-3293.

Pelikan, M. & Lobo, F. (2000). Parameter-less genetic algorithm: A worst-case time and space complexity analysis. *GECCO-2000 Proceedings of the Genetic and Evolutionary Computation Conference*, 370.

Reed, P., Minsker, B. S., & Goldberg, D. E. (2000). Designing a competent simple genetic algorithm for search and optimization. *Water Resources Research 36*(1), 3731–3741.

Reed, P., Minsker, B. S., Valocchi, A. J., & Goldberg, D. E. (2000). Efficient use of a genetic algorithm for long-term groundwater monitoring design. *Proceedings of the Thirteenth International Conference on Computational Methods in Water Resources.*

Rothlauf, F. & Goldberg, D. E. (2000). Pruefernumbers and genetic algorithms: A lesson how the locality of an encoding can harm the performance of GAs. *Proceedings of the 2000 Parallel Problem Solving from Nature Conference,* 395–404.

Rothlauf, F., Goldberg, D. E., & Heinzl, A. (2000). Bad codings and the utility of well-designed genetic algorithms. *GECCO-2000 Proceedings of the Genetic and Evolutionary Computation Conference*, 355–362.

Smalley, J. B., Minsker, B. S., & Goldberg, D. E. (2000). Risk-based in situ bioremediation using a noisy genetic algorithm. *Water Resources Research, 36*(10), 3043–3052.

Albert, L. A. (2001). *Efficient genetic algorithms using discretization scheduling* (MS thesis). General Engineering, University of Illinois at Urbana-Champaign, Urbana, IL.

Albert, L. & Goldberg, D. E. (2001). Efficient evaluation relaxation under integrated fitness functions. *GECCO-2001: Proceedings of the Genetic and Evolutionary Computation Conference*, 751.

Albert, L. & Goldberg, D. E. (2001). Efficient evaluation relaxation under integrated fitness functions. *Smart Engineering System Design: Neural Networks, Fuzzy Logic, Evolutionary Programming, Data Mining and Complex Systems*, 165–170.

Butz, M. V. (2001). *Anticipatory learning classifier systems.* Boston, MA: Kluwer Academic Publishers.

Butz, M. V., Goldberg, D. E., & Stolzmann, W. (2001). Probability-enhanced predictions in the anticipatory classifier system. In Lanzi, P. L., Stolzmann, W., & S. W. Wilson (eds.),*Advances in Learning Classifier Systems* (pp. 37-51). Berlin: Springer-Verlag.

Butz, M. V., Kovacs, T., Lanzi, P. L., & Wilson, S. W. (2001). How XCS evolves accurate classifiers. *GECCO-2001: Proceedings of the Genetic and Evolutionary Computation Conference*, 927–934.

Butz, M. V., & Pelikan, M. (2001). Analyzing the evolutionary pressure in XCS. *GECCO-2001: Proceedings of the Genetic and Evolutionary Computation Conference*, 935–943.

Espinoza, F., Minsker, B. S., & Goldberg, D. E. (2001). Optimal settings for a hybrid genetic algorithm applied to a groundwater remediation problem. *Bridging the Gap: Meeting the World's Water and Environmental Resources Challenges: Proceedings of the World Water and Environmental Resources Congress* [CD-ROM].

Espinoza, F. P., Minsker, B. S., & Goldberg, D. E. (2001). A self adaptive hybrid genetic algorithm. *GECCO-2001: Proceedings of the Genetic and Evolutionary Computation Conference*, 759.

Goldberg, D. E., & Sastry, K. (2001). A practical schema theorem for genetic algorithm design and tuning. *GECCO-2001: Proceedings of the Genetic and Evolutionary Computation Conference*, 328–335.

Goldberg, D. E., Sastry, K., & Latoza, T. (2001). On the supply of building blocks. *GECCO-2001: Proceedings of the Genetic and Evolutionary Computation Conference*, 336–342.

Gopalakrishnan, G., Minsker, B. S., & Goldberg, D. E. (2001). Optimal sampling in a noisy genetic algorithm for risk-based remediation design. *Bridging the Gap: Meeting the World's Water and Environmental Resources Challenges, Proceedings of the World Water and Environmental Resources Congress* [CD-ROM].

Gottlieb, J., Julstrom, B. A., Raidl, G. R., & Rothlauf, F. (2001). Prüfer numbers: A poor representation for spanning trees for evolutionary search. *GECCO-2001: Proceedings of the Genetic and Evolutionary Computation Conference*, 343–350.

Kosorukoff, A., Mittenthal, J., & Goldberg, D. (2001). Modeling of evolution of signaling networks in living cells by evolutionary computation. *GECCO-2001: Proceedings of the Genetic and Evolutionary Computation Conference*, 905.

Pelikan, M., & Goldberg, D. E. (2001). Escaping hierarchical traps with competent genetic algorithms. *GECCO-2001: Proceedings of the Genetic and Evolutionary Computation Conference*, 511–519.

Pelikan, M., & Goldberg, D. E. (2001). Hierarchical Bayesian optimization algorithm = Bayesian optimization algorithm + niching + local structures. *GECCO-2001 Workshop Proceedings*, 217–221.

Pelikan, M., Goldberg, D. E., Sastry, K. (2001). Bayesian optimization algorithm, decision graphs, and Occam's razor. *GECCO-2001: Proceedings of the Genetic and Evolutionary Computation Conference*, 519–526.

Parthasarathy, P. V., Goldberg, D. E., & Burns, S. A. (2001). Tackling multimodal problems in hybrid genetic algorithms. *GECCO-2001: Proceedings of the Genetic and Evolutionary Computation Conference*, 775.

Parthasarathy, P. V., Goldberg, D. E., & Burns, S. A. (2001). Tackling multimodal problems in hybrid genetic algorithms. *Smart Engineering System Design: Neural Networks, Fuzzy Logic, Evolutionary Programming, Data Mining and Complex Systems*, 123–128.

Reed, P., Minsker, B. S., & Goldberg, D. E. (2001). Designing a New Elitist Nondominated Sorted Genetic Algorithm for a Multiobjective Long Term Groundwater Monitoring Application. *GECCO-2001: Proceedings of the Genetic and Evolutionary Computation Conference*, 1454.

Reed, P. M., Minsker, B. S., & Goldberg, D. E. (2001). A multiobjective approach to cost effective long-term groundwater monitoring using an Elitist Nondominated Sorted Genetic Algorithm with historical data. *Journal of Hydroinformatics, 3*(2), 71–90.

Reed, P., Minsker, B. S., & Goldberg, D. E. (2001). The practitioner's role in competent search and optimization using genetic algorithms. *Bridging the Gap: Meeting the World's Water and Environmental Resources Challenges: Proceedings of the World Water and Environmental Resources Congress*, 129-134.

Rothlauf, F., Goldberg, D. E., & Heinzl, A. (2001). On the debate concerning evolutionary search using Prüfer numbers. *GECCO-2001 Workshop Proceedings, 262–267.*

Sastry, K. (2001). *Evaluation-relaxation schemes for genetic and evolutionary algorithms* (MS Thesis). General Engineering, University of Illinois at Urbana-Champaign, Urbana, IL.

Sastry, K. (2001). Efficient cluster optimization using extended compact genetic algorithm with seeded population. *GECCO-2001 Workshop Proceedings, 217–221.*

Sastry, K., & Goldberg, D. E. (2001). Modeling tournament selection with replacement using apparent added noise. *Smart Engineering System Design: Neural Networks, Fuzzy Logic, Evolutionary Programming, Data Mining and Complex Systems*, 129-134.

Sastry, K., & Goldberg, D. E. (2001). Modeling tournament selection with replacement using apparent added noise. *GECCO-2001: Proceedings of the Genetic and Evolutionary Computation Conference*, 781.

Sastry, K. Goldberg, D. E., & Pelikan, D. E. (2001). Don't evaluate, inherit. *GECCO-2001: Proceedings of the Genetic and Evolutionary Computation Conference*, 551–558.

Sinha, A., & Goldberg, D. E. (2001). Verification and extension of the theory of global-local hybrids. *GECCO-2001: Proceedings of the Genetic and Evolutionary Computation Conference*, 592–598.

Srivastava, R. P., & Goldberg, D. E. (2001). Verification of the theory of genetic algorithm continuation. *GECCO-2001: Proceedings of the Genetic and Evolutionary Computation Conference*, 623–630.

Tsutsui, S., & Goldberg, D. E. (2001). Search space boundary extension method in real-coded genetic algorithms. *Information Sciences, 133,*(3-4), 229–247.

Tsutsui, S., Goldberg, D. E., & Sastry, K. (2001). Simplex crossover and linkage learning in real-coded GAs. *GECCO-2001: Proceedings of the Genetic and Evolutionary Computation Conference*, 785.

Tsutsui, S., Goldberg, D. E., & Sastry, K. (2001). Linkage learning in real-coded GAs with simplex crossover. *Proceedings of the Fifth International Conference on Artificial Evolution*, 51–58.

Tsutsui, S., Pelikan, M., & Goldberg, D. E. (2001). Evolutionary algorithm using marginal histogram models in continuous domain, *GECCO-2001 Workshop Proceedings*, 230–233.

Tsutsui, S., Pelikan, M., & Goldberg, D. E. (2001). Probabilistic model-building genetic algorithms using marginal histograms in continuous domains, *Proceedings of KES'2001: The Fifth International Conference on Knowledge-Based Intelligent Information Engineering Systems & Allied Technologies*, 112-121.

van Hoyweghen, C., & Goldberg, D. E. (2001). Building block superiority, multimodality, and synchronization problems. *GECCO-2001: Proceedings of the Genetic and Evolutionary Computation Conference*, 694–701.

11

Albert, L. A. & Goldberg, D. E. (2002). Efficient discretization scheduling in multiple dimensions. *GECCO-2002: Proceedings of the Genetic and Evolutionary Computation Conference, GECCO-2002: Proceedings of the Genetic and Evolutionary Computation Conference,* 271–278.

Arst, R., Minsker, B. S., & Goldberg, D. E. (2002). Comparing advanced genetic algorithms and simple genetic algorithms for groundwater management. *ASCE Environmental & Waste Resources Institute (EWRI).*

Babbar, M., Minsker, B. S., & Goldberg, D. E. (2002). A multiscale island injection genetic algorithm for optimal groundwater remediation design. *ASCE Environmental & Waste Resources Institute (EWRI).*

Butz, M. V. (2002). *Anticipatory learning classifier systems.* Boston, MA: Kluwer Academic Publishers.

Butz, M. V., Goldberg, D. E., & Stolzmann, W. (2002). The anticipatory classifier system and genetic generalization. *Natural Computing, 1*(4), 427-467.

Butz, M. V., & Goldberg, D. E. (2002). Generalized state values in an anticipatory learning classifier system. *From Animals to Animats 7: The Seventh International Conference on the Simulation of Adaptive Behavior.*

Chen, J.-H., Goldberg, D. E., & Sastry, K. (2002). Fitness inheritance in multi-objective optimization. *GECCO-2002: Proceedings of the Genetic and Evolutionary Computation Conference,* 319–326.

Chen, Y.-P., & Goldberg, D. E. (2002). Introducing start expression genes to the linkage learning genetic algorithm. *Parallel Problem Solving from Nature, VII,* 351-360.

Goldberg, D. E. (2002). *The design of innovation: Lessons from and for competent genetic algorithms.* Boston, MA: Kluwer Academic Publishers.

Gopalakrishnan, G., Minsker, B. S., & Goldberg, D. E. (2003). Optimal sampling in a noisy genetic algorithm for risk-based remediation design. *Journal of Hydroinformatics, 5*(1), 11–26.

Khan, N., Goldberg, D. E., & Sastry, K. (2002). Multiple-objective Bayesian optimization algorithm. *GECCO-2002: Proceedings of the Genetic and Evolutionary Computation Conference,* 684.

Knjazew, D. (2002). *OmeGA: A competent genetic algorithm for solving permutation and scheduling problems.* Boston, MA: Kluwer Academic Publishers.

Kosorukoff, A., & Goldberg, D. E. (2002). Evolutionary computation as a form of organization. *GECCO-2002: Proceedings of the Genetic and Evolutionary Computation Conference,* 965–972. (Best MPP paper award)

Parthasarathy, P. V. (2002). *Hybrid genetic algorithm for sizing of members in fully-stressed design of frame structures* (MS thesis). General Engineering, University of Illinois at Urbana-Champaign, Urbana, IL.

Pelikan, M. (2002). *Bayesian optimization algorithm: From single level to hierarchy (PhD dissertation).* Computer Science, University of Illinois at Urbana-Champaign, Urbana, IL.

Pelikan, M., Goldberg, D. E., & Lobo, F. (2002). A survey of optimization by building and using probabilistic models. *Computational Optimization and Applications, 21*(5), 5–20

Pelikan, M., Goldberg, D. E., & Tsutsui, S. (2002). Combining the strengths of Bayesian optimization algorithm and adaptive evolutionary strategies. *GECCO-2002: Proceedings of the Genetic and Evolutionary Computation Conference*, 512-519.

Pelikan, M., Sastry, K., & Goldberg, D.E. (2002) Scalability of the Bayesian optimization algorithm. *International Journal of Approximate Reasoning, 31*(3), 221–258.

Rothlauf, F. (2002). *Towards a theory of representations for genetic and evolutionary algorithms* (Doctoral thesis). Management Information Science, University of Bayreuth, Germany.

Rothlauf, F. (2002). *Representations for genetic and evolutionary algorithms.* Heidelberg, Germany: Physica-Verlag.

Rothlauf, F., Goldberg, D. E., & Heinzl, A. (2002). Network random keys–a tree representation scheme for genetic and evolutionary algorithms. *Evolutionary Computation, 10*(1), 75–97.

Sastry, K. & Goldberg, D. E. (2002). Genetic algorithms, efficiency enhancement, and deciding well with differing fitness bias. *GECCO-2002: Proceedings of the Genetic and Evolutionary Computation Conference*, 536–543.

Sastry, K. & Goldberg, D. E. (2002). Genetic algorithms, efficiency enhancement, and deciding well with differing fitness variances. *GECCO-2002: Proceedings of the Genetic and Evolutionary Computation Conference*, 528–535.

Singh, A. & Goldberg, D. E., Chen, Y.-P. (2002). Modifying linkage learning genetic algorithm for difficult non-stationary problems. *GECCO-2002: Proceedings of the Genetic and Evolutionary Computation Conference*, 699.

Srivastava, R. P. (2002). *Time continuation in genetic algorithms* (MS thesis). General Engineering, University of Illinois at Urbana-Champaign, Urbana, IL.

Tsutsui, S., & Goldberg, D. E. (2001). Simplex crossover and linkage identification: Single-stage evolution versus multi-stage evolution. *IEEE Congress on Evolutionary Computation 2002*, 974–979.

Van Hoyweghen, C., Goldberg, D. E., Naudts, B. (2002). From twoMax to the Ising model: Easy and hard symmetrical problems. *GECCO-2002: Proceedings of the Genetic and Evolutionary Computation Conference*, 626–633.

Van Hoyweghen, C., Naudts, B., & Goldberg, D. E. (2002). Spin-flip symmetry and synchronization. *Evolutionary Computation, 10*(4), 317–344.

Espinoza, F., Minsker, B. S., & Goldberg, D. E. (2003). Local search issues for the application of a self-adaptive hybrid genetic algorithm in groundwater remediation design.

Khan, N. (2003). *Bayesian optimization algorithm for multiobjective and hierarchically difficult problems* (MS thesis). General Engineering, University of Illinois at Urbana-Champaign, Urbana, IL.

Singh, A., Minsker, B. S., Goldberg, D. E. (2003). Combining reliability and Pareto optimality—An approach using stochastic multi-objective genetic algorithms. *ASCE Environmental & Waste Resources Institute (EWRI)*.

Singh, A., Minsker, B. S., Goldberg, D. E. (2003). Combining reliability and Pareto optimality—An approach using stochastic multi-objective genetic algorithms. *ASCE Environmental & Waste Resources Institute (EWRI)*.

Sinha, A. (2003). *Designing efficient genetic and evolutionary algorithm hybrids* (MS thesis). General Engineering, University of Illinois at Urbana-Champaign, Urbana, IL.

5 Interactions and Transitions

This section lists meeting participation, presentations, and transitions.

5.1 Meeting participation and presentation

All conference papers above represent presentations by Professor Goldberg and his students. Professor Goldberg gave dozens of tutorials, keynote talks, tutorials, and other invited seminars and lectures across the country and around the world.

5.2 Transitions

ISGEC. The PI was founding chair (1999) for the International Society for Genetic and Evolutionary Computation, which sponsors the regular series of GECCO conferences. The 2003 conference was held in Chicago with 520 attendees.

Kluwer Book Series. The PI is the consulting editor of a book series called the *Kluwer Series on Genetic Algorithms and Evolutionary Computation* (http://www.wkap.nl/prod/s/GENA).

GA course online. PI's GA course (GE 485) is offered online via the web this semester. It is one of the most popular web courses offered by the University of Illinois and the National Technological University. See the website at http://online.cen.uiuc.edu/webcourses/ge485/index.html.

Innovation course online. The PI's book, *The Design of Innovation,* has been turned into an eight-lecture online short course (http://online.engr.uiuc.edu/shortcourses/innovation/index.html).

Cooling system design with Caterpillar. The PI has completed a GA-based cooling system design package for Caterpillar.

Telecommunications problems. A visitor from Germany, Franz Rothlauf (University of Bayreuth) and the PI worked on solving some telecommunications problems using random keys codings.

hBOA patent pending. The hierarchical Bayesian optimization algorithm is the subject of a patent filing by the University of Illinois. One company has already expressed interest in licensing hBOA, and negotiations are underway.

Technology transfer laboratory established. A rapidly growing software company in Israel established a research laboratory in Champaign to help transfer the competence and efficiency

results of this project to their commericial software in mobile telecommunications. Results to date have been promising.

Knjazew to SAP in Germany. Dimitri Knjazew who developed the OmeGA code has taken a position with SAP in Germany where he is adapting these techniques to SAP's product line.

Srivastava to Jump Trading. Ravi Srivastava took a job following his MS with a financial firm, Jump Trading, to apply GAs to solving problems in financial portfolio and trading analysis.

Spinoffs in IT and organizational theory. GA research is leading to spinoffs of the work in the creation of collaborative IT systems using HBGAs and interactive GAs. GA-style theory is leading to new quantitative models in organizational theory using elementary tradeoffs (ETOs).

GA-based companies. There are growing numbers of GA-based companies in the US and around the world using technology affected by my work. Optimax, a GA-based scheduling software firm has an integrated product for factory scheduling; Optimax recently sold out for close to $60M to i2, another scheduling software firm. The Schema Company in Israel has a GA-based software for the telecommunications business. Silicon Biology in Minnesota has developed a data-entry software package with GA-based technology at its core. Engineous Systems iSIGHT software use GAs as a base optimization engine. A center in England has spawned a company called Optimal Solutions that uses structured messy genetic algorithms to optimize water pipelines.

6 New discoveries, inventions, or patent disclosures

hBOA patent. A US patent filing was made in January 2002 for the hierarchical Bayesian optimization algorithm by the University of Illinois.

DSM GA and GAs for DSMs. Design structure matrices have formed the basis of a new competent GA and GAs have been used to cluster DSMs. Both sides of the DSM invention (GAs for DSMs and DSMs for GAs) have been disclosed to the university and a patent filing is anticipated.

7 Honors and Awards

This section lists awards for the PI and his students.

7.1 PI honors and awards

JSD Professor. Professor Goldberg was named Jerry S. Dobrovolny Distinguished Professor of Entrepreneurial Engineering in May 2003.

ISGEC Senior Fellow. Professor Goldberg was named part of the inaugural class of Senior Fellows for the International Society for Genetic and Evolutionary Computation in July 2003.

Other major awards. In 1997, Professor Goldberg was named a Gambrinus Fellow during his sabbatical to Dortmund, Germany. In 1996, Professor Goldberg received the ASEE Wickenden Award for best paper in the *Journal of Engineering Education* for his paper entitled "Change in

Engineering Education: One Myth, Two Scenarios, and Three Foci." In 1995, Professor Goldberg was named an Associate of the Center for Advanced Study (CAS) at the University of Illinois. In 1985, Professor Goldberg was named an NSF Presidential Young Investigator.

7.2 Student honors and awards

2002 CSE fellowship to Sastry. Kumara Sastry won a 2002–2003 Computational Science and Engineering fellowship at the University of Illinois at Urbana-Champaign for his work applying genetic programming to material science.

2003 CSE fellowship to Butz. Martin Butz won a 2003–2004 Computational Science and Engineering fellowship at the University of Illinois at Urbana-Champaign for his work in anticipation and genetics-based machine learning.

GECCO-2002 best paper. Alex Kosorukoff won best paper in methodology, philosophy, and pedagogy (MPP) at the 2002 GECCO for his paper "Evolutionary Computation as a Form of Organization" (with D. E. Goldberg).

GECCO-2003 best paper. Martin Butz and Kumara Sastry won best paper in classifiers systems (LCS) at the GECCO-2003 for their paper "Tournament Selection: Stable Fitness Pressure in XCS" with D. E. Goldberg).

GECCO-2002 best graduate student workshop paper. Y.-P. Chen took honors as best graduate student workshop paper at GECCO-2002 for his work on extending the linkage learning GA.